Metacognition Learning (2013) 8:79–102
DOI 10.1007/s11409-013-9096-5

# First- and second-order metacognitive judgments of semantic memory reports: The influence of personality traits and cognitive styles

**Sandra Buratti · Carl Martin Allwood · Sabina Kleitman**

Received: 30 August 2012 / Accepted: 28 January 2013 /
Published online: 13 February 2013
© Springer Science+Business Media New York 2013

**Abstract** In learning contexts, people need to make realistic confidence judgments about their memory performance. The present study investigated whether second-order judgments of first-order confidence judgments could help people improve their confidence judgments of semantic memory information. Furthermore, we assessed whether different personality and cognitive style constructs help explain differences in this ability. Participants answered 40 general knowledge questions and rated how confident they were that they had answered each question correctly. They were then asked to adjust the confidence judgments they believed to be most unrealistic, thus making second-order judgments of their first-order judgments. As a group, the participants did not increase the realism of their confidence judgments, but they did significantly increase their confidence for correct items. Furthermore, participants scoring high on an openness composite were more likely to display higher confidence after both the first- and second-order judgments. Moreover, participants scoring high on the openness and the extraversion composites were more likely to display higher levels of overconfidence after both the first- and second-order judgments. In general, however, personality and cognitive style factors showed only a weak relationship with the ability to modify the most unrealistic confidence judgments. Finally, the results showed no evidence that personality and cognitive style supported first- and second-order judgments differently.

**Keywords** Metacognitive accuracy · Second-order judgments · Personality · Cognitive style · Confidence judgments · Calibration

S. Buratti (✉)
Department of Psychology, University of Gothenburg, Box 500, 40530 Göteborg, Sweden
e-mail: sandra.buratti@psy.gu.se

C. M. Allwood
Department of Psychology, University of Gothenburg, Göteborg, Sweden

S. Kleitman
School of Psychology, University of Sydney, Sydney, Australia

## Introduction

Metacognition is commonly defined as "any knowledge or cognitive activity that takes as its object, or regulates, any aspect of any cognitive enterprise" (Flavell et al. 1993, p. 150), and such judgments are important in learning contexts. For example, people may evaluate the correctness of their retrieved memories of studied material by making explicit or implicit confidence judgments about the correctness of the retrieved memory (Koriat and Goldsmith 1996), and obviously, confidence judgments for learnt material need to be as accurate (realistic) as possible. Confidence judgments can be deficient in different ways. For example, people can be *overconfident*, with a confidence level indicating that they are more confident than they are correct, or they can be *underconfident* and correct to a greater extent than their level of confidence indicates.

Confidence judgments can be seen as *first-order* judgments, and the questioning and adjustment of the accuracy of these judgments can be seen as *second-order* judgments. People often show overconfidence in their first-order judgments of their memory performance for semantic knowledge (e.g. Griffin and Brenner 2004; Kleitman and Stankov 2001; Lichtenstein et al. 1982; McClelland and Bolger 1994), and overconfidence can have serious implications for students, such as when they plan study time or take an exam. To illustrate, overconfidence might lead students to prematurely stop studying or to hand in an incorrect answer on an exam. Thus, it is important to identify ways people can improve the accuracy of their confidence judgments.

There are a number of important contexts where the performance of successful *second-order* confidence judgments are helpful and can help to adjust earlier unrealistic first-order judgments. For example, in a context outside of the world of education, an eyewitness to a crime might be asked to reconsider their level of confidence for a certain memory report and it is important that the level of confidence a witness feel for a certain memory report is accurate since jurors often judge a witness' credibility based on their confidence level (e.g. Cutler et al. 1988). The goal to improve the accuracy of second-order judgments may be especially relevant in educational contexts in which confidence judgments are important for optimizing learning outcomes (Wang et al. 1990). For example, students when reviewing their answers on an examination, need to make accurate second-order judgments about their confidence concerning different answers on the test. If the students find that they have been unrealistically confident about a certain answer this will guide them to reconsider and possibly change the answer to one they feel more confident about (Kleitman and Costa in press). Likewise, when students review their theses or essays, second-order judgments will guide them concerning which sections of their texts they earlier might have felt unrealistically confident about and therefore now need to consider for rewriting.

The present study investigated participants' ability to improve the accuracy of first-order confidence judgments about the correctness of retrieved semantic knowledge. Moreover, previous research has identified associations (albeit sometimes weak) between certain personality traits and cognitive styles and confidence judgment accuracy (Dahl et al. 2010; Kleitman and Stankov 2007; Pallier et al. 2002; Schaefer et al. 2004; Want and Kleitman 2006); thus, another goal was to investigate which traits and styles share meaningful relationships with successful second-order confidence judgments. Furthermore, we examined if these traits and styles are the same as those associated with accurate first-order confidence judgments.

Previous research has investigated the use of second-order judgments in different contexts (for a review see, Serra and England 2012). Some of these studies have investigated to which extent people can improve metacognitive judgments by performing second-order

judgments (confidence judgments: Buratti and Allwood 2012a, b; confidence intervals: Cesarini et al. 2006; judgments of learning: Dunlosky et al. 2005; global judgments: Miller and Geraci 2011). However, to our knowledge, no previous research has analyzed the ability to improve confidence judgment accuracy regarding the correctness of semantic memory reports.

## First- and second-order metacognitive judgments

Miller and Geraci (2011) found that low-performing students who showed overconfidence (i.e. were more confident than correct) in predicting their total exam performance were more accurate in their second-order judgment of their performance than high-performing students. Likewise, Dunlosky et al. (2005) found evidence of successful second-order assessments about judgments regarding learning (JoLs) of paired associates. In two studies, Buratti and Allwood (2012a, b) found that people can regulate first-order confidence judgments of their reports for an episodic memory event. These first-order confidence judgments are performed in what we below call the *Confidence task*, that is, when a person makes confidence judgments about the correctness of each answer they give. In one of these studies (2012a), the participants were asked to *exclude* the confidence judgments they believed were most unrealistic that is, the confidence level was not judged to be in line with each answer's correctness. Participants were able to make their confidence judgments more accurate, although the effect was small. In the other study (2012b), participants were asked to *adjust* the confidence judgments they believed were unrealistic to make them more accurate (we call the task of identifying confidence judgments that need improvement and then carrying out these adjustments the *Adjustment task*). As in the 2012a study, participants improved the accuracy of their confidence judgments, although this effect also was small. Thus, people seem to be able to improve confidence judgment accuracy for episodic memory reports as well as other types of metacognitive judgments.

There are different ways to measure confidence accuracy. In the present study, just as in Buratti and Allwood (2012a, b), we looked at the participants' ability to improve the so called *absolute* accuracy of their confidence judgments. Absolute accuracy refers to the direct relation between the participants' confidence judgments and the correctness of their memory-reports (summarized over many items). In measures of the *relative* accuracy of confidence judgments, the participants' ability to discriminate, or separate, correct and incorrect items by means of the level of their confidence judgments is considered. However, although we primarily look at the absolute accuracy of the participants confidence judgments, our study also indirectly relates to the relative accuracy of their confidence judgments in that we also perform our analyses separately for correct from incorrect answers. Thus, when the participants are successful in improving their confidence accuracy for corrects answers it means that they increase the level of their confidence closer to 100 % (the ideal level, *Absolutely sure the answer is correct*), thereby making their confidence judgment more extreme and thus also improving the relative accuracy of their confidence judgments. Similarly, for incorrect answers, when the participants are successful in improving their confidence accuracy it means that they decrease the level of their confidence, closer to 0 % (the ideal level, *Absolutely sure the answer is incorrect*).

Second-order judgments can be seen as meta-metacognitive because they focus on first-order metacognitive judgments. Figure 1 shows our conceptualization of how meta-metacognition relates to metacognition and cognitive processes and how the two different parts of the Adjustment task relate to meta-metacognition. We suggest that the identification part of the Adjustment task (where the confidence judgments that need adjusting are

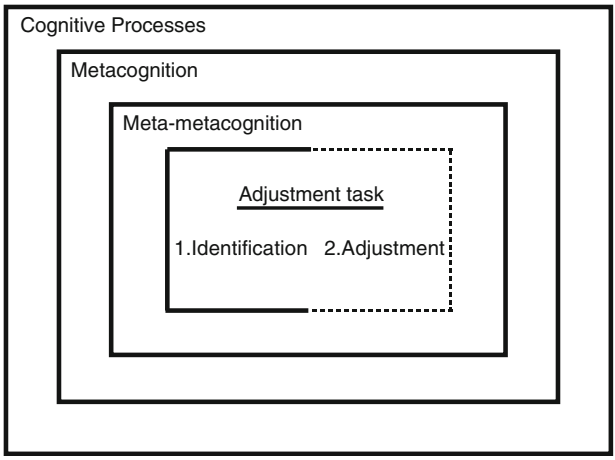


**Fig. 1** The conceptual relation between the two parts of the Adjustment task, meta-metacognition and metacognition

identified) should be seen as a meta-metacognitive task because it takes a metacognitive product as its target. For this reason, the identification part of the Adjustment task is of special interest for analysis. In contrast, the adjustment part is not necessarily a meta-metacognitive task. Here the participant may simply perform the identified confidence judgment again, although now under the assumption that it has been identified as a confidence judgment needing revision. Thus, the adjustment part of the Adjustment task is marked with a dashed line in Figure 1 because its status as meta-metacognition or meta-cognition is controversial. In addition, the extent to which meta-metacognitive and meta-cognitive processes differ in nature from other cognitive processes remains an issue for future research. However, people may consider somewhat different things in meta-metacognitive tasks compared with metacognitive tasks; for example, they may be more willing to consider arguments against their selected answer to the knowledge. This model is in line with Nelson and Narens' model (1990, 1994) of metacognition in which metacognitive monitoring and control can occur on several levels, that is metacognitive monitoring and control of metacognitive processes as well as of cognitive processes but these authors never used the term meta-metacognition to describe the control and monitoring of meta-cognitive processes.

## Personality and cognitive styles as predictors of first- and second-order confidence judgments

In light of the research on second-order judgments, it is of interest to ask: What personality and cognitive style factors predict the success of first- and second-order judgments? Worth noting is that different personality and cognitive styles factors may be important predictors for first-order and second-order judgments, respectively. We next review the relevant literature and describe our expectations with respect to factor analytic results based on measures we predicted would be relevant to a study of first- and second-order judgments. In this context, we also present our expectations with respect to the effect on first- and second-order judgments of some of the included personality and cognitive styles.

*Extraversion and narcissism* Research on first-order confidence judgments has shown small, yet statistically stigmificant, associations between extraversion and overconfidence (Dahl et al. 2010; Pallier et al. 2002; Schaefer et al. 2004). A partial explanation may be that extraversion is associated with individuals who are active, sociably competent, and optimistic (Costa and McCrae 1988). Thus, people high in extraversion might feel they are competent in general and as a consequence show high confidence regardless of whether the answer is correct or wrong.

Moreover, research has found a positive correlation between narcissism, i.e. people who have a grandiose sense of self-importance (Ronningstam 2005), and overconfidence (Campbell et al. 2004). Because a high level of narcissism is correlated with illusions of greater intelligence (Gabriel et al. 1994), people high in narcissism are likely more confident about their answers without being more correct than people on average.

Also worth noting is that studies have reported a correlation between narcissism and extraversion (e.g. Vazire et al. 2008). For this reason, we expected that narcissism and extraversion would define the same factor and that this factor would be related to participant performance in the Confidence and Adjustment tasks. In other words, we expected that the Extraversion/Narcissism factor would show a positive correlation with bias in the Confidence task; with respect to the Adjustment task, we expected that people high in the Extraversion/Narcissism factor would be less willing to doubt the accuracy of their first-order judgments and thus show negative correlations with the number of adjustments made. For this reason, and given their tendency to high bias, we expected poorer performance in the Adjustment task overall for people high in the Extraversion/Narcissism factor compared to people low in this factor.

*Openness, need for cognition, and close-mindedness* The Openness (or Intellect) personality dimension is defined by being open to possibilities and different solutions and having an active imagination, an appreciation of inner feelings and insights, and intellectual curiosity (e.g. Goldberg 1992). Previous research has found Openness to share positive correlations with both proportion of correct answers and the level of confidence (Dahl et al. 2010; Kleitman 2008; Schaefer et al. 2004). In general, people high on this personality trait tend to perform better on different cognitive (intelligence) tests, and this might be an underlying common cause for its association with confidence judgments. We expected that people high on this trait would be more confident and more accurate in their responses on a semantic memory task.

Similarly, Need for Cognition (NFC) concerns people who enjoy solving complex problems and problematizing issues (Cacioppo et al. 1984). Some researchers have found a small positive association between NFC and overconfidence (Wolfe and Grosch 1990), but others have not (Jonsson and Allwood 2003). Moreover, NFC has been associated with the Openness personality dimension (e.g. Sadowski and Cogburn 1997). In contrast, tendencies towards mental rigidity and Need for Closure (NFCl; Kruglanski 1989), especially the Close-minded facet, have negative associations with the Openness personality trait (Kleitman 2008). The NFCl construct is defined as "the desire for a definite answer on some topic, any answer as opposed to confusion and ambiguity" (Kruglanski 1989, p. 14). Feelings of discomfort with ambiguity act as a motivation for this construct, and it includes four other broad tendencies: preference for predictability, preference for order, decisiveness, and closed-mindedness. Thus, we predicted that these constructs would define the same factor, with Openness and NFC defining it with positive loading and the Closed-minded aspect of the NFCl scale with negative. Consistent with earlier findings, we expected that people high in this factor (labeled "Openness") would have a higher proportion of correctly answered knowledge questions and higher confidence ratings on both the Confidence

(Kleitman and Stankov 2007; Pallier et al. 2002) and Adjustment tasks. We also expected that people high in the Openness factor would perform a lower number of adjustments because previous research has found that people high in Openness tend to be more correct and thus to have higher confidence, which may lead to a general tendency not to doubt their previous judgments (Schaefer et al. 2004). The more unstable results for the NFC scale might function to dilute this prediction.

*Self-doubt* We employed several measures to assess self-doubt tendencies: the Judgmental Self-doubt Scale (capturing perceptions of self-doubt with respect to one's ability to make decisions; Mirels et al. 2002), Self-doubt subscale (capturing self-doubts with respect to one's ability in general; Oleson et al. 2000), and Clance Imposter Phenomenon Scale (assesses subjective fears of evaluation and inability to repeat previous success; Clance 1985). We expected that responses on these scales would converge into a Self-doubt factor indicating self-doubt and hesitance in decision tendencies. We also expected that the Emotional stability personality dimension (tendencies to feel relaxed and not worrying) and several aspects of the NFC construct (Roets and Van Hiel 2007; Webster and Kruglanski 1994), i.e. the tendencies to feel uncomfortable with ambiguity and to feel decisive and thus make decisions quickly to avoid uncertainty, would load meaningfully on this Self-doubt factor.

Earlier studies have found a negative correlation between feelings of self-doubt and confidence for different judgment tasks (Mirels et al. 2002; Want and Kleitman 2006). Moreover, when it comes to making successful second-order metacognitive judgments, the person making the judgments may need to be able to genuinely question their judgment ability. Persons high in self-doubt may be more willing to go back and change first-order judgments that they suspect are not accurate. In general, having a greater general tendency to doubt one's first-order judgments might lead to being open to reevaluating and processing the first-order judgment more deeply, or at least differently. Thus, we expected that the Self-doubt factor would show negative associations with on-task confidence judgments (for both the Confidence and the Adjustment tasks) and that the Self-doubt factor would correlate positively with a willingness to identify and then adjust confidence judgments determined to need it.

*Conscientiousness and agreeableness* Conscientiousness is defined as a tendency to act dutifully and show self-discipline. Some previous research has reported positive small-size correlations between overconfidence and conscientiousness (Dahl et al. 2010; Schaefer et al. 2004). However, other researchers have not replicated these correlations (e.g. Kleitman 2008; Kleitman and Costa in press; see also Stankov 1999 and Stankov and Kleitman 2008 for reviews). For this reason, we held no expectations regarding the overall relations between Conscientiousness and first- and second-order judgments. In line with previous findings, we expected, however, that Conscientiousness would show meaningful positive correlations and define the same factor with the Need for Order and Predictability facets of the NFCl scale (Kruglanski 1989). These three measures were thus expected to constitute a separate factor.

Similarly, we held no direct predictions with respect to the Agreeableness personality dimension (defined as being social and considerate of other people). A full Big Five personality measure was, however, included in the study to preserve the validity of the responses, and it was included in the overall exploratory factor analyses.

## Aims of the current study

The aims of this study were, first, to improve our understanding of second-order judgments of semantic memory reports (general knowledge questions) and, second, to investigate

which personality and cognitive style facets relate to confidence level after first- and second-order confidence judgments of semantic memory reports. For the second aim, we first determined the factorial structure of the personality measures and cognitive styles employed in this study using an exploratory factor analysis. Then we used the derived factor scores in a series of hierarchical multiple regressions to determine the predictors of the first- and second-order confidence judgments.

## Method

### Participants

In total, 151 persons participated (46 men and 105 women) with a mean age of 22.4 years (SD=4.2, range 16–45). A majority of the participants ($N$=124) were first year psychology students participating as part of their course requirements at an Australian university. The other participants were students recruited on campus who were reimbursed with 15 AUD per hour for their participation. The study was approved by the human ethics committee at the university where the study was carried out (14021).

### Procedure

First the participants listened to a 10-min presentation regarding the concept *confidence realism*. They were given examples of people being on average underconfident (less confident than correct); perfectly realistic (no difference in confidence and portion correct); and overconfident (more confident than correct). Two examples for each (mis)calibration instance were provided. Participants were also shown the confidence scale that they were to use later and were given a thorough example of how to calculate the level of realism in confidence judgments by comparing the average level of confidence to the proportion correct. To measure how well they understood the concept, participants then answered 10 questions concerning the concept *confidence realism*.

Next, during the *Confidence task*, the participants answered 40 general knowledge questions on different topics such as world geography, literature, and biology. After each question, they rated how sure they were that their answer was correct on a confidence scale ranging from 0 %, "I am absolutely sure that my answer is incorrect," to 100 %, "I am absolutely sure that my answer is correct."

The participants were asked to answer all questions, and if they did not know the answer, they were to guess. They were informed that the person with the highest proportion correct would be rewarded with a movie ticket.

After completing the *Confidence* task, the participants next performed the *Adjustment task*. In this task, the participants were presented with their answers and their confidence judgments of their answers. They were told that their aim for the Adjustment task was to be as realistic as possible on average in their confidence judgments. To achieve this aim, they were asked to modify the earlier confidence judgments they believed were unrealistic by providing new ones on the new confidence scale that followed each question. They were allowed to modify as few or as many as they wanted but were recommended to change at least 20 confidence judgments. The reason why we made this recommendation was that we wanted to prompt the participants to make a serious adjustment attempt but still have the freedom to not adjust that many confidence judgments ($n$=20) if they really felt that they had already been highly realistic in their first-order judgments. They were told that the

person with the best confidence realism after this task would be rewarded with a movie ticket.

The participants then answered the questionnaires regarding personality and cognitive styles. The order of the questionnaires was counterbalanced to avoid order effects. However, to avoid influence on the self-doubt measures of the perceived achievement on the general knowledge task, questionnaires measuring self-doubt were never placed directly after the Adjustment task. The overall testing time was about 1.5 h, depending on the individual response time.

## Measures of metacognitive accuracy

To measure how well confidence judgments match the level of correctness of the memory reports, several measures can be used (Yates 1994). One commonly used measure of confidence accuracy is the *bias* measure, also called *over-/underconfidence*, which is calculated by taking the average confidence level of the memory reports and subtracting the proportion correct. A negative value indicates underconfidence while zero indicates perfect realism and a positive value indicates overconfidence. Thus, the measure provides information regarding the direction of the bias.

In the context of the present study, when people are asked to make second-order judgments with the aim of increasing the realism in their confidence judgments, their aim is to get as close to zero as possible on the bias measure regardless of whether they are under- or overconfident to begin with. Using bias as a measure could then cancel out improvements made by participants on average. Therefore, we used the bias measure as a descriptive measure but used the *absolute bias* measure for analyses of improvements in confidence realism. The absolute bias measure is calculated by taking the absolute value of each participant's bias score.

The first portion of the Adjustment part, in which confidence judgments to be adjusted are identified, clearly can be seen as involving a second-order judgment; thus, this part was of special interest for further analysis. We therefore analyzed participant success in identifying the confidence judgments that needed adjustment. In this analysis, we evaluated correct and incorrect answers separately. The further away a confidence judgment is from its normatively ideal point (e.g. 100 % for correct answers and 0 % for incorrect answers), the more it is in need of adjustment.

To analyze success in identifying confidence judgments needing adjustment, for the correct answers and for each participant, we summarized the difference between 100 % and the level of each specific identified confidence judgment. For the incorrect answers, the sum of each specific confidence judgment was calculated. Thus, the more the confidence for an answer deviated from the normatively ideal value, the more it was in need of adjustment; accordingly, the more it contributed to this sum. The sums, one for correct and one for incorrect answers, respectively, thus reflect the accumulated confidence deviating from the normatively ideal value. For these measurements to be meaningful, the accumulated sums must also be related to how much confidence the participant could correct optimally (that is, taking into consideration how good their realism was to start with). Thus, for correct and incorrect answers separately, we took the ratio between each of the just-described sums for the identified confidence

judgments and the corresponding sums for all the confidence judgments made in the Confidence task. In this case, the closer this ratio is to 1 for both correct and incorrect answers, the greater the success of the participant in identifying the confidence judgments in need of adjustment. Thus, the formula for the *correct answers* identified to be adjusted is (1):

$$\frac{\sum_{t=1}^{T}(1-\mathrm{ConfCIA_t})}{\sum_{f=1}^{F}(1-\mathrm{ConfCA_f})} \quad (1)$$

where $\mathrm{ConfCIA_t}$=confidence for a correct identified answer t, $\mathrm{ConfCA_f}$=confidence for a correct answer f, T=total number of correct identified answers, and F=total number of correct answers.

Correspondingly, the computations for the *incorrect answers* identified to be adjusted were made according to the following formula (2):

$$\frac{\sum_{u=1}^{U}(\mathrm{ConfIIA_u})}{\sum_{g=1}^{G}(\mathrm{ConfIA_g})} \quad (2)$$

where $\mathrm{ConfIIA_u}$=confidence for an incorrect identified answer u, $\mathrm{ConfIA_g}$=confidence for an incorrect answer g, U=total number of incorrect identified answers, and G=total number of incorrect answers.

## Measures of personality and cognitive styles

*Goldberg's Big Five Markers* (GBFM) are representations of the Goldberg (1992) markers for the big-five personality dimensions: *Extraversion* ("Start conversations" and "Don't mind being the center of attention"), *Agreeableness* (e.g. "Am interested in people" and "Take time out for others"), *Conscientiousness* ("Pay attention to details" and "Get chores done right away"), *Intellect* ("Spend time reflecting on things" and "Am quick to understand things"), and *Emotional stability* ("Am relaxed most of the time" and "Seldom feel blue"). Each subscale contains 10 items that were rated on a 5-point scale (1=Very inaccurate, 5=Very accurate).

*Judgmental Self-doubt Scale (Mirels et al. 2002)* This scale consists of 19 items that assess feelings of self-doubt with respect to one's ability to make decisions. Two example statements are, "I have difficulty making decisions" and "I have a tendency to change my mind according to the last opinion I hear." The answers were given on a 6-point scale (−3=Strongly disagree, 3=Strongly agree).

*Self-doubt subscale (Oleson et al. 2000)* This subscale contains 9 items that assess feelings of self-doubt and was taken from the Subjective Overachievement Scale presented by the same authors (examples: "More often than not I feel unsure of my abilities" and "I sometimes find myself wondering if I have the ability to succeed at important activities"). The answers were given on a 6-point scale (1=Disagree very much, 6=Agree very much).

*Clance imposter phenomenon scale (Clance 1985)* The scale consists of 20 items that measure feelings of self-doubt in one's performance, feelings of being a fraud or an imposter ("I can give the impression that I'm more competent than I really am" and "I avoid evaluations if possible and have a dread of others evaluating me"). The items were rated on a 5-point scale (1=Not at all true, 5=Very true).

*Narcissistic personality inventory (Raskin and Terry 1988)* The scale consists of 40 items that assess the degree of narcissism in non-clinical samples ("I have a natural talent for influencing people" and "Modesty doesn't become me"). The participants choose one statement with which they most agree out of two choices.

*Need for cognition (Cacioppo et al. 1984)* The scale consists of 18 items that assess a person's tendency to engage in and enjoy cognitively demanding tasks. Two example statements are, "I would prefer complex to simple problems" and "The notion of thinking abstractly is appealing to me." The answers were given on a 9-point scale (−4=Very strong disagreement, 4=Very strong agreement).

*Need for closure scale (Webster and Kruglanski 1994)* The questionnaire consists of 41 items divided into 5 subscales: *Order* (10 items), which assesses a person's preference for order and structure ("I think that having clear rules and order at work is essential for success" and "I hate to change my plans at the last minute"); *Predictability* (8 items), which assesses a person's need predict sure outcomes ("I don't like to go into a situation without knowing what I can expect from it" and "I don't like to be with people who are capable of unexpected actions"); *Decisiveness* (6 items, we used the revised Decisiveness subscale by Roets and Van Hiel 2007), which assesses a person's assertiveness ("When I have made a decision, I feel relieved" and "When I am confronted with a problem, I'm dying to reach a solution very quickly"); *Ambiguity* (9 items), which assesses a person's discomfort with ambiguity ("I don't like situations that are uncertain" and "I like to know what people are thinking all the time"); and finally, *Close-mindedness* (8 items), which measures a person's lack of openness to new ideas and solutions ("I dislike questions which could be answered in many different ways" and "I do not usually consult many different opinions before forming my own view"). The items were rated on 6-point scale (1=Strongly disagree, 6=Strongly agree).

## Results

All data analysis was done using SPSS *v*19 software. On the test concerning the understanding of the concept *confidence realism*, 97 % of the participants provided at least 8 (out of 10) correct answers, indicating that the majority of participants showed a good understanding of the concept described to them by the researcher. Table 1 shows the proportion of correct answers and the level of confidence and metacognitive measures for the Confidence and Adjustment tasks. Reliabilities were calculated by using the method employed by Stankov and Crawford (1996). Split-half reliabilities were calculated by correlating the odd-numbered items with the even-numbered items. The correlations were corrected using the Spearman-Brown formula.

On average, participants adjusted 6.7 ($SD$=5.4; range 0–22) confidence judgments in the Adjustment task. There was a significant negative correlation between the bias level in the Confidence task and how many items that were adjusted ($r$=− 0.21, $p$=0.011). Thus, the higher a participant's bias value, the lower the number of adjusted items.

**Table 1** Descriptive statistics for proportion correct, confidence and realism measures across tasks and correlations with factor scores ($n$ =151)

| | | | | | | Correlations | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | Reliability | Age | Sex | S | O | C | E/N | A |
| *Confidence task* | | | | | | | | | | |
| Proportion correct | 0.519 | 0.117 | – | 0.208* | −0.166* | −0.197* | 0.329** | −0.073 | −0.088 | −0.088 |
| Confidence | 0.551 | 0.123 | 0.89 | 0.263** | –0.299** | –0.258** | 0.437** | –0.007 | 0.053 | −0.105 |
| Bias | 0.032 | 0.073 | 0.60 | 0.110 | −0.240** | −0.121 | 0.211** | 0.106 | 0.232** | −0.036 |
| Absolute bias | 0.063 | 0.049 | 0.42 | 0.014 | −0.104 | 0.058 | 0.093 | −0.020 | 0.032 | −0.182* |
| Confidence for correct items | 0.877 | 0.085 | – | 0.174* | −0.296** | −0.209* | 0.267** | 0.065 | 0.202* | −0.072 |
| Confidence for incorrect items | 0.204 | 0.088 | – | 0.203* | −0.264** | −0.159 | 0.382** | 0.039 | 0.131 | −0.072 |
| *Adjustment task* | | | | | | | | | | |
| Number of adjustments | 6.7 | 5.4 | – | −0.082 | 0.109 | 0.024 | −0.015 | 0.067 | −0.054 | 0.075 |
| Confidence | 0.560 | 0.120 | 0.86 | 0.253** | −0.279** | −0.262** | 0.434** | −0.016 | 0.037 | −0.093 |
| Bias | 0.041 | 0.065 | 0.49 | 0.090 | −0.215** | −0.128 | 0.206* | 0.102 | 0.225** | −0.013 |
| Absolute bias | 0.065 | 0.042 | 0.23 | −0.007 | −0.098 | 0.068 | 0.146 | −0.056 | 0.042 | −0.112 |
| Confidence for correct items | 0.900 | 0.077 | – | 0.151 | −0.199* | −0.204* | 0.239** | 0.093 | 0.197* | −0.014 |
| Confidence for incorrect items | 0.201 | 0.086 | – | 0.181* | −0.275** | −0.165* | 0.368** | −0.007 | 0.098 | −0.082 |

*S* Self-doubt, *O* Openness, *C* Conscientiousness, *E/N* Extraversion/Narcissism, *A* Agreeableness

* $p$<.05. ** $p$<.01

To investigate whether the participants on average managed to increase the realism in their confidence judgments, a repeated measures ANOVA was calculated between the absolute bias values calculated for the Confidence and Adjustment tasks. The results showed that the participants did not manage to significantly increase their confidence realism (in terms of both over- and underconfidence). However, they did manage to increase their confidence significantly for correct items ($F(1,150)=34.642$, $p<0.001$, partial $\eta^2=0.12$), but there was no significant difference between the confidence level for incorrect items after the Confidence task and the Adjustment task. In this context, we note that there was a negative correlation between the confidence level for correct items in the Confidence task and the number of items that were adjusted ($r=-0.32$, $p<0.001$). Thus, people with lower confidence for correct items in the Confidence task chose to adjust more items in the Adjustment task. No corresponding correlation was found for incorrect items.

### Identification of the most unrealistic items

If participants were able to identify the most unrealistic items, then the average confidence level for *correct* items not chosen to be modified should be higher than the confidence level for the correct items chosen to be modified. Results showed that the confidence level for correct items not chosen to be modified ($M=0.93$, $SD=0.09$) was higher than the confidence level for the correct items chosen to be modified ($M=0.66$, $SD=0.17$; $t(113)=15.15$, $p<0.001$). The same procedure was used to analyze the confidence level for not-chosen and chosen incorrect items. For incorrect items, the average confidence level for not-chosen items should ideally be lower than the average confidence level for chosen items. The results showed that the items not chosen to be modified ($M=0.13$, $SD=0.09$) had a significantly lower confidence level than the items chosen to be modified ($M=0.47$, $SD=0.20$; $t(109)=-16.91$, $p<0.001$). Thus, both for correct and incorrect items, the participants tended to select suitable items when they identified items for adjustment. A correlation was also calculated to investigate how successful the participants were in choosing the more biased items for adjustment. The correlation between the bias level for each item and whether an item was chosen for modification or not was calculated for each individual and the average correlation across participants was .42[1] ($n=122$). This correlation indicates, in line with the results reported above, that the participants tended to choose items with a higher degree of bias.

### Factor analysis

Table 2 shows the descriptive statistics and reliability measures for the different personality and cognitive styles questionnaires.

The student sample in this study had a mean value of 65.11 on the Clance Imposter Phenomenon Scale, which is slightly higher than other similar samples, which often score about 50 (King and Cooley 1995; Ross et al. 2001). This outcome indicates that the participants in this sample have somewhat higher feelings of being a fraud or an imposter compared with other student samples. The mean value of the Narcissistic personality inventory was close to that reported by Raskin and Terry (1988) for a student sample. The results with respect to NFCl and personality measures were similar to those from previous

[1] It should be noted that there are some constraints to this correlation since the number of adjustments varied a lot between participants. Also correlations could not be calculated for participants that did not choose to do any adjustments.

**Table 2** Descriptive statistics, reliability estimates (Cronbach's Alpha) and correlations for the questionnaires employed ($n$=150)

| | *M* | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Personality variables* | | | | | | | | | | | | | | | | | |
| 1. GBFM: Extraversion | 33.89 | 7.62 | (0.88) | 0.18* | 0.00 | 0.15 | 0.06 | −0.32** | −0.33** | −0.28** | 0.54** | 0.04 | 0.06 | −0.26** | −0.08 | 0.04 | −0.01 |
| 2. GBFM: Agreeableness | 40.73 | 5.11 | | (0.77) | 0.10 | −0.01 | 0.08 | 0.07 | 0.01 | 0.01 | −0.22** | 0.11 | 0.15 | −0.06 | −0.09 | −0.01 | −0.23** |
| 3. GBFM: Conscientiousness | 31.27 | 6.18 | | | (0.75) | 0.01 | 0.03 | −0.08 | −0.08 | −0.04 | 0.01 | 0.04 | 0.59** | 0.29** | −0.02 | 0.08 | 0.01 |
| 4. GBFM: Emotional stability | 29.58 | 7.29 | | | | (0.83) | 0.08 | −0.54** | −0.54** | −0.42** | 0.02 | 0.09 | −0.24** | −0.34** | −0.31** | −0.49** | −0.01 |
| 5. GBFM: Intellect | 37.61 | 5.44 | | | | | (0.75) | −0.34** | −0.13 | −0.03 | 0.15 | 0.567** | −0.05 | −0.19* | −0.16* | −0.06 | −0.37** |
| 6. Judgmental Self-doubt Scale | 0.11 | 1.28 | | | | | | (0.94) | 0.62** | 0.45** | −0.27** | −0.34** | 0.15 | 0.267** | 0.34** | 0.29** | 0.10 |
| 7. Self−doubt Scale | 30.36 | 6.46 | | | | | | | (0.80) | 0.72** | −0.27** | −0.23** | 0.20* | 0.32** | 0.41** | 0.39** | 0.03 |
| 8. Clance Imposter Phenomenon | 65.11 | 12.35 | | | | | | | | (0.88) | −0.18* | −0.13 | 0.22** | 0.34** | 0.33** | 0.36** | −0.06 |
| 9. Narcissistic Personality Inventory | 13.03 | 7.15 | | | | | | | | | (0.86) | 0.08 | 0.04 | −0.03 | 0.01 | 0.22** | 0.14 |
| *Cognitive styles* | | | | | | | | | | | | | | | | | |
| 10. Need for Cognition | 63.64 | 11.05 | | | | | | | | | | (0.90) | −0.07 | −0.25** | −0.22** | −0.10 | −0.46** |
| 11. NFCl: Order | 3.82 | 0.85 | | | | | | | | | | | (0.81) | 0.48** | 0.22** | 0.41** | −0.09 |
| 12. NFCl: Predictability | 3.49 | 0.88 | | | | | | | | | | | | (0.79) | 0.47** | 0.51** | 0.26** |
| 13. NFCl: Decisiveness | 3.60 | 0.91 | | | | | | | | | | | | | (0.76) | 0.59** | 0.27** |
| 14. NFCl: Ambiguity | 4.25 | 0.75 | | | | | | | | | | | | | | (0.73) | 0.16 |
| 15. NFCl: Close-mindedness | 2.87 | 0.61 | | | | | | | | | | | | | | | (0.54) |

Cronbach's alpha estimates are given in brackets in the diagonal. *GBFM* Goldberg's Big Five Markers, *NFCl* Need for Closure Scale

* $p$<.05. ** $p$<.01. *** $p$<.001

research with Australian undergraduates using the same or similar measures (Kleitman 2008). As the diagonal of Table 2 indicates, almost all measures showed high reliability. The only exception was the close-mindedness scale, which had an alpha value of only 0.52.

In addition, Table 2 summarizes the correlation coefficients between self-report measures employed in this study. The pattern of relationships between variables is consistent with previous research and is best summarized by the Exploratory Factor Analysis (EFA) described next.

To determine the factorial structure of the measured personality and cognitive style variables, an EFA was conducted (principal component with Promax rotation). Table 3 presents the results.

The latent root criteria suggested a five factor solution, and it explained 71.6 % of the total variance. The resulting factors made good theoretical sense, and some replicated previous findings (see below). Thus, this solution was accepted and the interpretation of the factors is as follows:

Factor 1. *Self-doubt.* The Self-doubt scale, Clance Imposter Phenomenon, and GBFM: Emotional stability negatively loaded highest on this factor. Also loading on this factor was the Judgmental Self-doubt Scale and the two NFCl facets: Ambiguity and Decisiveness. NFCl: Predictability had a small loading as well as Conscientiousness, which loaded negatively. Clearly, as hypothesized, this factor represents a tendency towards self-doubt.

Factor 2. *Openness.* As predicted, Openness/Intellect and the NFC scale both loaded positively on this factor with the Close-mindedness facet of the NFC scale loading negatively. This factor replicates earlier findings (Kleitman 2008)[2] and indexes tendencies to enjoy novelty and cognitive engagement and to be open to different solutions to problems.

Factor 3. *Conscientiousness.* Replicating earlier results (Kleitman 2008), this factor was defined by the Conscientiousness personality dimension and the two NFCl facets: Order and Predictability. This factor indexes tendencies to plan, to strive for order in life, and to be meticulous.

Factor 4. *Extraversion/Narcissism.* As expected, Extraversion and the Narcissistic Personality Inventory loaded highest on this factor. The Ambiguity facet of the NFCl scale also loaded, although with a smaller loading. This factor indexes tendencies to be vocal, sociable, active, self-focused, and authoritative.

Factor 5. *Agreeableness.* Agreeableness and Extraversion personality dimensions loaded on this factor. The Close-mindedness facet of the NFCl loaded negatively, although with a small loading. This factor indexes people who care about what other people think about them and are sociable, friendly, and considerate.

The bottom part of Table 3 displays the correlations between factors. Generally, the factors shared no, or small, correlations. The Self-doubt factor correlated negatively with the Openness and the Agreeableness factors and positively with the Conscientiousness factor. The Conscientiousness factor also correlated positively with the Extraversion/Narcissism factor and negatively with the Agreeableness factor. The Extraversion/Narcissism factor correlated positively with the Agreeableness factor.

The correlations between the factors (calculated using relevant factor scores) and the measures for the Confidence and Adjustment tasks are presented in Table 1. The sizes of the

[2] The results with respect to the Need for Cognition scale are, although novel (for example, it was not included in the Kleitman 2008 study), consistent with the nature of this factor.

**Table 3** Results of an EFA (PC with Promax Rotation) on cognitive styles and personality measures ($n$=151)

| | F1 Self-doubt | F2 Openness | F3 Conscientiousness | F4 Extraversion/ Narcissism | F5 Agreeableness | $h^2$ |
|---|---|---|---|---|---|---|
| Self-doubt scale | 0.854 | | | | | 0.773 |
| GBFM: Emotional Stability | −0.805 | | | | | 0.575 |
| Clance Imposter Phenomenon | 0.795 | | | | | 0.635 |
| NFCl: Ambiguity | 0.716 | | | 0.397 | | 0.750 |
| Judgemental Self-Doubt scale | 0.671 | | | | | 0.698 |
| NFCl: Decisiveness | 0.601 | | | | | 0.551 |
| GBFM: Openness/ Intellect | | 0.871 | | | | 0.719 |
| Need for Cognition | | 0.852 | | | | 0.727 |
| NFCl: Close-mindedness | | −0.640 | | | −0.387 | 0.648 |
| GBFM: Conscientiousness | −0.333 | | 0.960 | | | 0.719 |
| NFCl: Order | | | 0.844 | | | 0.802 |
| NFCl: Predictability | 0.316 | | 0.560 | | | 0.708 |
| GBFM: Extraversion | | | | 0.894 | 0.424 | 0.846 |
| Narcisstic Personality Inventory | | | | 0.826 | | 0.769 |
| GBFM: Agreeableness | | | | | 0.902 | 0.745 |
| *Factor intercorrelations* | | | | | | |
| F1: Self-doubt | 1 | −0.284** | 0.304** | −0.103 | −0.166* | |
| F2: Openness | | 1 | −0.096 | 0.039 | 0.191* | |
| F3: Conscientiousness | | | 1 | 0.179* | −0.185* | |
| F4: Extraversion/ Narcissism | | | | 1 | −0.210** | |
| F5: Agreeableness | | | | | 1 | |

*GBFM* Goldberg's Big Five Markers, *NFCl* Need for Closure Scale

* $p$<.05. ** $p$<.01

correlation coefficients vary and most are small, but some consistent patterns did emerge. Replicating previous findings (Kleitman 2008; Pallier et al. 2002; Schaefer et al. 2004), the Openness factor correlated most consistently with the dependent variables on the Confidence and Adjustment tasks (correlations range between 0.211 to 0.437, $p$<0.01). That is, Openness is moderately correlated with proportion correct and confidence level in both tasks, indicating that people high in Openness tend to have a larger proportion of correct answers but also tend to be even more confident in their answers on both tasks, which explains the significant correlation between bias and the Openness factor.

As predicted, the Self-doubt factor correlated most consistently with the variables of interest (correlations ranging from −0.197 to −0.262, $p$<0.01). There was a small, significant negative correlation with proportion correct and confidence both in the Confidence task and

the Adjustment task. There were also small but significant negative correlations with confidence for correct items in the Confidence task and both confidence for correct and incorrect items in the Adjustment task. This result replicates earlier findings in which confidence for different judgment tasks was negatively correlated with feelings of self-doubt (Mirels et al. 2002; Want and Kleitman 2006).

In line with previous research, the Narcissism/Extraversion factor correlated significantly with the bias measure and confidence for correct answers (Pallier et al. 2002; Campbell et al. 2004; Schaefer et al. 2004; Dahl et al. 2010). No consistent correlations existed between the Conscientiousness and Agreeableness factors and the different measures (although there was a small correlation between the Agreeableness factor and bias). Despite our expectations, *the number of items* adjusted in the Adjustment task did not correlate with *any* of the factors.

## Personality and cognitive styles factors and the capability of identifying items in need of adjustment

We next investigated whether correlations between how successful the participants were at identifying items in need of adjustment and the different personality and cognitive style factors. For this purpose, we first calculated two ratios, one for correct items and one for incorrect items, as described above, in which each ratio has a possible value range from 0 to 1. A value of 0 indicates that the participant did not identify any part of the confidence that deviated from the normatively ideal confidence value (100 for correct answers and 0 for incorrect answers) after the Confidence task. A value of 1 indicates that the participant, in the Adjustment task, identified all confidence that deviated from the ideal confidence value. Participants who had fewer than 5 unrealistic correct or incorrect items were excluded from the analyses.[3] The participants identified on average 46 % of the potential deviation from perfect realism for correct answers and on average 39 % of the potential deviation for incorrect answers. Each of these two ratios was then correlated with the personality and cognitive style factors. These analyses showed no significant correlations, and we found no quadratic relationships.

We note that the ratio for correct items was highly correlated with the ratio for incorrect items, $r=0.71$, $p<0.001$.[4] This finding indicates that participants who were successful in identifying correct items in need of adjustment also were successful at identifying incorrect items in need of adjustment. In addition, the *more* items the participants modified, the better they were at identifying items needing adjustment for both correct ($r=0.78$, $p<0.001$) and incorrect items ($r=0.82$, $p<0.001$). In spite of the fact that the participants neglected to identify many items in need of adjustment, further analyses showed that participants were succesful at not choosing items that were *not in need of adjustment*, since only 1 % of the correct items with "100 %" first-order confidence judgments and only 1 % of the incorrect items with "0 %" first-order confidence judgments were identified for adjustment.

## Personality and cognitive style factors and confidence in first- and second-order confidence judgments

To investigate which personality variables predict overall confidence levels for Confidence and Adjustment tasks, two hierarchical regression analyses were performed. Accordingly,

[3] 94 participants were left in the analysis for correct answers and 121 participants in the analysis for incorrect answers.

[4] This analysis was done only for the participants who had more than 5 unrealistic items when calculating both the correct and incorrect ratio; thus, these correlations consist of 75 participants.

**Table 4** Hierarchical multiple regression analyses predicting confidence level from personality and cognitive styles factor scores

| Predictor | Confidence level | | | |
|---|---|---|---|---|
| | Confidence task | | Adjustment task | |
| | $\Delta R^2$ | β | $\Delta R^2$ | β |
| Step 1 | 0.707*** | | 0.746*** | |
| Age | | 0.101* | | 0.084 |
| Gender | | −0.171*** | | −0.145*** |
| Proportion correct | | 0.769*** | | 0.807*** |
| Step 2 | 0.035** | | 0.031** | |
| F1: Self-doubt | | −0.052 | | −0.056 |
| F2: Openness | | 0.141** | | 0.130** |
| F3: Conscientiousness | | 0.057 | | 0.055 |
| F4: Extraversion/Narcissism | | 0.093* | | 0.081 |
| F5: Agreeableness | | 0.002 | | 0.008 |
| Total $R^2$ | 0.742*** | | 0.777*** | |

* $p<.05$. ** $p<.01$. *** $p<.001$

confidence levels for the Confidence and the Adjustment tasks were regressed on the five identified factors (in block 2), after controlling for the effects of age, sex, and accuracy of performance (entered as block 1). The results are summarized in Table 4.

Given the inclusion of accuracy of performance, the first block of variables predicted a massive 70.7 % and 74.6 % of variance in confidence levels for the Confidence and Adjustment tasks, respectively. Not surprisingly, the proportion of correctly answered items was the strongest predictor of confidence levels for both tasks. Females had significantly lower confidence levels compared to males, although the prediction was not very strong. Age contributed weakly toward the prediction of confidence levels for the Confidence task only; older students had somewhat higher levels of confidence. Beyond these predictions, personality variables added 3.5 % and 3.1 % of variance toward the prediction of confidence levels in both tasks, respectively ($p< 0.01$). Openness was a significant but weak predictor of confidence in both tasks, and Extraversion/Narcissism was a significant but weak predictor in the Adjustment task only.

Hierarchical regressions were also performed with confidence for correct and for incorrect items as dependent variables in each of the Confidence and Adjustment tasks, respectively, to explore if the factors predict confidence for these differently. In Step 1, confidence levels for correct and incorrect items were regressed on age and sex.[5] Then, in Step 2, the personality and cognitive style variables were added to the model. Table 5 summarizes the results.

Age and sex predicted a significant amount of variance in each of the four regression equations (11.7 %, 6.2 %, 11.0 %, and 10.8 %, respectively) with sex being a consistently negative predictor. Females held significantly lower levels of confidence for correct and incorrect items on both tasks. After controlling for age and sex, Extraversion/Narcissism

[5] We did not control for the accuracy of performance because such control will be redundant given that confidence levels were split on the basis of the correctly and incorrectly answered items.

**Table 5** Hierarchical multiple regression analyses predicting confidence level for correct and incorrect items from personality and cognitive styles factors scores

| Predictor | Confidence level for correct items | | | | Confidence level for incorrect items | | | |
|---|---|---|---|---|---|---|---|---|
| | Confidence task | | Adjustment task | | Confidence task | | Adjustment task | |
| | $\Delta R^2$ | β | $\Delta R^2$ | β | $\Delta R^2$ | β | $\Delta R^2$ | β |
| Step 1 | 0.117*** | | 0.062** | | 0.110*** | | 0.108*** | |
| Age | | 0.072* | | 0.150 | | 0.202** | | 0.179* |
| Gender | | −0.295*** | | −0.198* | | −0.263*** | | −0.274** |
| Step 2 | 0.082* | | 0.094* | | 0.012** | | 0.097** | |
| F1: Self-doubt | | −0.120 | | −0.142 | | −0.039 | | −0.044 |
| F2: Openness | | 0.161 | | 0.146 | | 0.325*** | | 0.312*** |
| F3: Conscientiousness | | 0.086 | | 0.124 | | 0.039 | | 0.000 |
| F4: Extraversion/Narcissism | | 0.161* | | 0.160 | | 0.095 | | 0.062 |
| F5: Agreeableness | | 0.010 | | 0.046 | | −0.057 | | −0.077 |
| Total $R^2$ | 0.199*** | | 0.154*** | | 0.222*** | | 0.205*** | |

* $p<.05$. ** $p<.01$. *** $p<.001$

significantly predicted confidence level for correct items in the Confidence task but not in the Adjustment task. As in previous analyses, this prediction was not very strong. For incorrect items, Openness significantly predicted confidence level both in the Confidence task and in the Adjustment task.

## Discussion

In this study, we investigated participants' ability to successfully adjust their first-order confidence judgments of their semantic memory reports. In this context, we also took an individual difference approach and analyzed which personality and cognitive style factors underlie success in performing first- and second-order metacognitive judgments of semantic memory information.

As a group, the participants did not increase the realism in their confidence judgments in the Adjustment task. At least two reasons might explain why these results differ from those reported by Buratti and Allwood (2012b) in which participants successfully modified their confidence judgments in an episodic memory task, thus making their confidence judgments more accurate. First, semantic memory reports may differ from episodic reports, for example in providing less informative cues about the accuracy of the confidence judgments. Second, the level of accuracy of the confidence judgments in the present study may have been too good (i.e. too realistic) to allow ample room for improvement. Future studies should use items which render higher level of bias to investigate if participants are better at adjusting their realism when there is ample room for improvements.

In addition, the participants managed to increase the confidence for the correct items, which indicated that they thus partly were able to decrease the bias in their confidence judgments. Accordingly, they were more successful at handling the correct than the incorrect items. At a general level, it can be noted that this result is in line with the general tendency of confirmation bias (Nickerson 1998) because the adjustments for correct items involved increasing the confidence level of the adjusted items. This result is also of interest when compared with the results of Buratti and Allwood (2012b), who found that participants were better at handling the incorrect items. One possible reason for this discrepancy could be the difference in memory tasks used. When it comes to semantic memory situations, people often receive feedback on their performance, in contrast to episodic memory situations Perfect et al. (1993). Thus, if a person remembers a plausible answer to a general knowledge question, they likely will have a stronger feeling of familiarity about the answer than for an encoded episodic memory. In the present study, the participants therefore may have experienced a stronger feeling of familiarity about correct items than for incorrect items, and it is reasonable to think that they had encoded correct answers more often than incorrect answers. Subsequently, the participants would be more likely to increase their confidence for these items because of a feeling of familiarity. This feeling would in this situation serve as a cue for correct items but not for incorrect items because incorrect answers may not have had the same encoding possibilities. There might be further reasons for the difference in adjusting confidence judgments in episodic and semantic memory. For example, Perfect (2004) suggested that the making of confidence judgments involves use of different heuristics such as speed of retrieval, mental imagery and vividness. These heuristics might be more important for judgments of episodic memory than semantic memory and further studies should investigate whether these heuristic also are important for second-order judgments and if they could help explain the differences found between the result for second-order judgments of episodic and semantic memory.

However, the fact that the participants were able to increase their confidence significantly for correct items indicates that they were able to make succesful second-order judgments. This result is line with Dunlosky et al. (2005) who found that their participants were succesful at making second-order assessments of another type of metacognitive judgments, namely JoLs. The results in this study is also in line with the Miller and Geraci (2011) findings where student were found to be better at assessing the validity of their first-order judgments than they were at making their first-order judgments of their exam performance. A further conclusion from the result that the participants managed to improve their confidence accuracy for correct items is that they thereby contributed to improving their relative confidence accuracy. This, since they increased the level of the adjusted confidence judgments for the correct items and did not worsen the confidence accuracy for the incorrect items.

Moreover, further analysis showed that the participants, on average, had some ability to *identify* confidence judgments in need of adjustment, as compared to the confidence judgments not selected for adjustment. The participants identified on average 46 % of the potential deviation from perfect realism for correct answers and 39 % of the potential deviation for incorrect answers. Future research should investigate how people can be helped to identify more of the confidence judgments in need of adjustment. For example, first, they could be made more aware of the fact that most people need to identify more items that would profit from adjustment and second, they could be given more specific advice on what type of items to look for.

Another interesting result was that participants with higher bias level adjusted fewer items. Thus, participants that were less realistic in their first-order confidence judgments were also less able to see this flaw in their realism. In brief, participants that had a higher bias had two deficiencies, their confidence judgments lacked in realism and they seemed to be unaware of this flaw. Future research should further investigate the mechanisms relating these two deficiencies for this type of participants.

We next discuss the results from the personality and cognitive style factor analysis, which provided a basis for further analyses regarding first- and second-order judgments. As anticipated, the Self-doubt measures loaded together with the NFCl subscales Ambiguity and Decisiveness, defining the Self-doubt factor. Also, as anticipated, the Openness personality dimension and the NFC scale loaded on one factor, and the Close-mindedness scale loaded negatively on this factor as well. We also anticipated that the extraversion and the Narcissistic Personality Inventory would load together (e.g. Vazire et al. 2008). Of interest, the ambiguity facet of the NFCl also loaded on this factor, although the loading was small. Finally, two factors emerged that we had not hypothesized about, namely the Conscientiousness and the Agreeableness factors. The Conscientiousness factor consisted of the Conscientiousness personality dimension and two NFCl subscales, which capture related tendencies: need for order and predictability. The Agreeableness factor consisted of the agreeableness and extraversion subscales, and in addition, the close-mindedness subscale loaded negatively.

With respect to the relation between the personality and cognitive style factors and success in the second-order judgments, no correlation was found between participants' ability to identify the items needing adjustment and the personality and cognitive style factors. The reason could be that there were two floor effects. As noted above, the participants had a fairly good level of realism in their confidence judgments in the Confidence task, which left little room for further improvements in the Adjustment task. The second floor effect resulted from the low number of changes made. In general, the low number of changes indicated an absence of ability to identify items that would be beneficial

to adjust. By opting for a low number of changes, the participants decreased their chances of performing well in the Adjustment task.

Overall, our hierarchical regressions showed that the Openness factor significantly predicted a high confidence level for both the first- and second-order judgments and that the Narcissism/Extraversion significantly predicted a high confidence level for the Confidence task (first-order judgment) and nearly so for the Adjustment task. When the confidence level was investigated separately for correct and incorrect items, we found that Openness predicted the level of confidence for incorrect items but not for correct items in both first- and second-order judgments.

Moreover, the zero-order correlations showed that people high in Openness did not identify more items to be modified in the Adjustment task than other participants. In addition, after the Adjustment task, they were still overconfident about their incorrect answers while being realistic about their correct answers (i.e. not underconfident). The fact that people high in Openness were more confident about incorrect answers than were people low in Openness accords with our speculation that people high in Openness would be less likely to question their answers and thus less able to try to regulate their realism for incorrect answers. This concordance could possibly be explained by the tendency for participants who were high in Openness to believe in their intellectual ability and curiosity. Belief in their intellectual ability/curiosity may have hindered these participants from doubting the correctness of an answer, thus making them more confident about incorrect answers. This inference is further supported by the fact that in our study, the Openness factor was negatively correlated with the Self-doubt factor. In brief, although people high in openness may be more open to the world than other people, they may not be more open than other people to doubting their own judgments.

According to the hierarchical regression analyses, the Extraversion/Narcissism factor significantly predicted confidence in first-order judgments for correct items. Also, significant correlations showed that these people were biased in their first-order judgments and continued to be biased after their second-order judgments. This last result is in line with earlier research indicating that people high in extraversion (Pallier et al. 2002) and narcissism (Schaefer et al. 2004) tend to be more biased.

Surprisingly, none of the factors were linked to the *number of adjustments* made in the Adjustment task. Theoretically, the people high in Self-doubt and Conscientiousness should have been more likely to adjust more confidence judgments while people high in Narcissism and Extraversion should have been more likely to adjust fewer. However, there were no such tendencies in the results of this study.

A fair amount of research has been done that investigates the association between personality and metacognitive judgments (e.g. Pallier et al. 2002; Want and Kleitman 2006; Washburn et al. 2005). Overall, it seems that personality and cognitive style factors are only weakly linked to the making of first- and second-order confidence judgments. In line with our results, Dahl et al. (2010) found only a very low association between personality and metacognitive realism in first-order judgments for elderly people and semantic memory. However, there may be differences between semantic and episodic memory when it comes to first- and second-order confidence judgments. Based on the findings of the present study, and assuming that there is no important difference between first- and second-order confidence judgments, it is possible that personality variables such as openness might be important not just for confidence judgments but for example also when making JoLs and tip-of-the-tongue judgments. For example, it might be, as stated above, that people high in openness believe in their intellectual ability and therefore are more likely to give high level JoLs, irrespectively of whether they later can recall the item or not. It is also

possible that people high in openness are more likely to actually recall the item (have better knowledge or memory). This suggestion illustrates the need for further research on the influence of personality variables on metacognitive judgments. For example, since there might be differences between semantic and episodic memory, it may also be interesting to investigate the impact of personality and cognitive style factors while controlling for intelligence on metacognitive judgments of episodic memory.

Also, in regards to the regulation of confidence accuracy, future studies should attempt to get participants to engage even further in the Adjustment task since a lack of engagement in this task could be one reason why the improvement effect was so small in our study.

**Acknowledgment** This research was partially supported by a grant given to the second author from the Swedish Research Council (VR). Special thanks to the Australian Endeavour Awards for financially supporting the visit of the first author to the University of Sydney. Special thanks to Simon Jackson for excellent suggestions regarding the manuscript.

## References

Buratti, S., & Allwood, C. M. (2012a). The accuracy of meta-metacognitive judgments: Regulating the realism of confidence. *Cognitive Processing, 13*(3), 243–253. doi:10.1007/s10339-012-0440-5.

Buratti, S., & Allwood, C.M. (2012b). Improved realism of confidence for an episodic memory event. *Judgment and Decision Making*, *7*(5), 590–601. Retrieved from http://journal.sjdm.org/11/111121a/jdm111121a.html.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306–307. doi:10.1207/s15327752jpa4803_13.

Campbell, W. K., Goodie, A. S., & Foster, J. D. (2004). Narcissism, confidence, and risk attitude. *Journal of Behavioral Decision Making, 17*(4), 297–311. doi:10.1002/bdm.475.

Cesarini, D., Sandewall, Ö., & Johannesson, M. (2006). Confidence interval estimation task and the economics of overconfidence. *Journal of Economic Behavior & Organization, 61*, 453–470. doi:10.1016/j.jebo.2004.10.010.

Clance, P. R. (1985). *The Imposter Phenomenon: Overcoming the fear that haunts your success*. Atlanta: Peachtree.

Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO personality inventory. *Journal of Personality and Social Psychology, 54*(5), 853–863. doi:10.1037/0022-3514.54.5.853.

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*, 41–55. doi:10.1007/bf01064273.

Dahl, M., Allwood, C. M., Rennemark, M., & Hagberg, B. (2010). The relation between personality and the realism in confidence judgements in older adults. *European Journal of Ageing, 7*(4), 283–291. doi:10.1007/s10433-010-0164-2.

Dunlosky, J., Serra, M. J., Matvey, G., & Rawson, K. A. (2005). Second-order judgments about judgments of learning. *The Journal of General Psychology, 132*(4), 335–346. doi:10.3200/GENP.132.4.335-346.

Flavell, J.H., Miller, S.A., & Miller, P.H. (1993). *Cognitive development* (3rd ed.) Englewood Cliffs, NJ: Prentice-Hall.

Gabriel, M. T., Critelli, J. W., & Ee, J. S. (1994). Narcissistic illusions in self-evaluations of intelligence and attractiveness. *Journal of Personality, 62*(1), 143–155. doi:10.1111/j.1467-6494.1994.tb00798.x.

Goldberg, L. R. (1992). The development of markers for the Big Five structure. *Psychological Assessment, 4*, 26–42.

Griffin, D., & Brenner, L. (2004). In D. J. Koehler & N. Harvey (Eds.), *Perspectives on probability judgment calibration*. Malden: Blackwell Publishing.

Jonsson, A.-C., & Allwood, C. M. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. *Personality and Individual Differences, 34*(4), 559–574. doi:10.1016/S0191-8869(02)00028-4.

King, J. E., & Cooley, E. L. (1995). Achievement orientation and the impostor phenomenon among college students. *Contemporary Educational Psychology, 20*(3), 304–312. doi:10.1006/ceps.1995.1019.

Kleitman, S. (2008). *Metacognition in the rationality debate. Self-confidence and its calibration*. Germany: VDM Verlag Dr. Mueller.
Kleitman, S., & Costa, D. (in press). The role of a novel formative assessment tool (Stats-mIQ) and individual differences in real-life academic performance. *Learning and Individual Differences.*
Kleitman, S., & Stankov, L. (2001). Ecological and person-oriented aspects of metacognitive processes in test-taking. *Applied Cognitive Psychology, 15*(3), 321–341. doi:10.1002/acp.705.
Kleitman, S., & Stankov, L. (2007). Self-confidence and metacognitive processes. *Learning and Individual Differences Journal, 17*(2), 161–173.
Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490–517. Vazire.
Kruglanski, A. W. (1989). *Lay epistemics and human knowledge*. New York: Plenum Press.
Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of subjective probabilities: The state of the art up to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.
McClelland, A. G. R., & Bolger, F. (1994). In G. Wright & P. Ayton (Eds.), *The calibration of subjective probability: Theories and models 1980–94*. New York: John Wiley and Sons.
Miller, T. M., & Geraci, L. (2011). Unskilled but aware: Reinterpreting overconfidence in low-performing students. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 502–506. doi:10.1037/a0021802.
Mirels, H. L., Greblo, P., & Dean, J. B. (2002). Judgmental self-doubt: Beliefs about one's judgmental prowess. *Personality and Individual Differences, 33*(5), 741–758. doi:10.1016/S0191-8869(01)00189-1.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 1–45). New York: Academic.
Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge: MIT Press.
Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220. doi:10.1037/1089-2680.2.2.175.
Oleson, K. C., Poehlmann, K. M., Yost, J. H., Lynch, M. E., & Arkin, R. M. (2000). Subjective overachievement: Individual differences in self-doubt and concern with performance. *Journal of Personality, 68*(3), 491–524. doi:10.1111/1467-6494.00104.
Pallier, G., Wilkinson, R., Danthir, V., Kleitman, S., Knezevic, G., Stankov, L., et al. (2002). The role of individual differences in the accuracy of confidence judgments. *The Journal of General Psychology, 129*(3), 257–299. doi:10.1080/00221300209602099.
Perfect, T. J. (2004). The role of self-rated ability in the accuracy of confidence judgements in eyewitness memory and general knowledge. *Applied Cognitive Psychology, 18*(2), 157–168. doi:10.1002/acp.952.
Perfect, T. J., Watson, E. L., & Wagstaff, G. F. (1993). Accuracy of confidence ratings associated with general knowledge and eyewitness memory. *Journal of Applied Psychology, 78*(1), 144–147. doi:10.1037/0021-9010.78.1.144.
Raskin, R., & Terry, H. (1988). A principal components analysis of the narcissistic personality inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54*, 890–902.
Roets, A., & Van Hiel, A. (2007). Separating ability from need: Clarifying the dimensional structure of the Need for Closure scale. *Personality and Social Psychology Bulletin, 33*(2), 266–280. doi:10.1177/0146167206294744.
Ronningstam, E. F. (2005). *Identifying and understanding the narcissistic personality*. New York: Oxford University Press.
Ross, S. R., Stewart, J., Mugge, M., & Fultz, B. (2001). The imposter phenomenon, achievement dispositions, and the five factor model. *Personality and Individual Differences, 31*(8), 1347–1355. doi:10.1016/S0191-8869(00)00228-2.
Sadowski, C. J., & Cogburn, H. E. (1997). Need for cognition in the Big-Five factor structure. *Journal of Psychology: Interdisciplinary and Applied, 131*(3), 307–312.
Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, W. K. (2004). Overconfidence and the Big Five. *Journal of Research in Personality, 38*(5), 473–480. doi:10.1016/j.jrp.2003.09.010.
Serra, M. J., & England, B. D. (2012). Magnitude and accuracy differences between judgements of remembering and forgetting. *The Quarterly Journal of Experimental Psychology, 65*(11), 2231–2257. doi:10.1080/17470218.2012.685081.
Stankov, L. (1999). Mining on the "no man's land" between intelligence and personality. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 315–337). Washington, DC: American Psychological Association.
Stankov, L., & Crawford, J. D. (1996). Confidence judgments in studies of individual differences. *Personality and Individual Differences, 21*(6), 971–986. doi:10.1016/S0191-8869(96)00130-4.

Stankov, L., & Kleitman, S. (2008). Processes on the borderline between cognitive abilities and personality: Confidence and its realism. In G. J. Boyle, G. Matthews, & D. Saklofske (Eds.), *The SAGE handbook of personality theory and testing* (Vol. 1, pp. 541–555). Thousand Oaks: Sage Publications.

Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2008). Portrait of a narcissist: Manifestations of narcissism in physical appearance. *Journal of Research in Personality, 42*, 1439–1447. doi:10.1016/j.jrp.2008.06.007.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *The Journal of Educational Research, 84*(1), 30–43.

Want, J., & Kleitman, S. (2006). Imposter phenomenon and self-handicapping: Links with parenting styles and self-confidence. *Personality and Individual Differences, 40*(5), 961–971. doi:10.1016/j.paid.2005.10.005.

Washburn, D. A., Smith, J. D., & Taglialatela, L. A. (2005). Individual differences in Metacognitive responsiveness: Cognitive and personality correlates. *The Journal of General Psychology, 132*(4), 446–461. doi:10.3200/GENP.132.4.446-461.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*(6), 1049–1062. doi:10.1037/0022-3514.67.6.1049.

Wolfe, R. N., & Grosch, J. W. (1990). Personality correlates of confidence in one's decisions. *Journal of Personality, 58*, 515–534. doi:10.1111/j.1467-6494.1990.tb00241.x.

Yates, J. F. (1994). Subjective probability accuracy analysis. In G. A. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 381–410). Oxford: John Wiley and Sons.